# ANOMALY PREDICTION BASED ON LSTM AND AUTOENCODERS USING FEDERATED LEARNING IN FINANCIAL TRANSACTIONS

**Synopsis submitted to IIMT University in partial fulfillment of the requirement for the award of the degree of**

**DOCTOR OF PHILOSOPHY**
**IN**
**COMPUTER SCIENCE**


**By**
**Gaurav Kumar**
**(Enrollment No - IIMTU/21233046)**

**Under the guidance of**
**Dr. Pushpendra Kumar Verma**
**(Associate Professor, SoCSA, IIMTU)**


**IIMT UNIVERSITY, MEERUT**

**SCHOOL OF COMPUTER SCIENCE & APPLICATIONS**
**IIMT UNIVERSITY, MEERUT (UP)**
**FEBRUARY- 2025**

Gaurav Kumar

Dean
School of Computer Sci
IIMT University, Meerut (UP)

## Table of Contents


1. Abstract ..... 1
2. Introduction ..... 1
3. The problem ..... 2
4. Gap in the body of knowledge and motivation of the research ..... 3
5. Aim and research questions ..... 4
6. Objectives of the study ..... 5
7. Theoretical background ..... 5
8. Scope of study ..... 9
9. Literature review ..... 10
10. Research methodology ..... 13
11. Discussion ..... 15
12. Conclusion and Future Scope ..... 17
13. References ..... 18

## 1. Abstract

In the financial sector, predictive models are essential for assessing risk and supporting decision-making processes. But sensitive nature of financial data creates significant issue for data sharing and collaborative model development. Privacy-preserving federated learning algorithms designed to train financial risk prediction models while maintaining data privacy. It provides a detailed insight of Federated Machine Learning, including existing distributed architectures and several privacy-preserving methods. Privacy mechanism ensures data confidentiality during the collaborative instruction process. The way Federated Learning represents an arising area of future that promises to revolutionize AI and Machine Learning by enabling secure and cooperative analytics. The perception presented in this research underscore the importance of privacy-preserving methods in fostering innovation and trust in financial analytics.

## 2. Introduction

The advancement of Artificial Intelligence from the recent decades has led to notable advancements in fields such as robotics, computer intelligence, and applications used in management. However, a considerable concern is the preservation of data privacy [1]. With immeasurable amount of information generated daily, ensuring data privacy has become increasingly important [2]. Data breaches involving both publicly available and private corporate data have heightened the need for stringent data privacy measures. Utilizing data isolated in "data islands" while nurture strict privacy standards is key for enhanced data security as shown in figure 1.

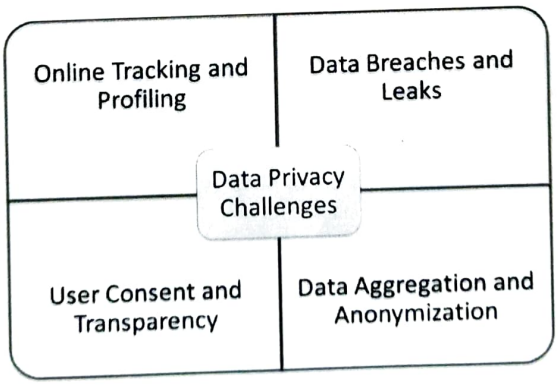


Fig 1: Data Privacy Challenges

Misuse of personal information can lead to user discomfort and reluctance to share personal details [3]. In companies and industries, protecting data from leaks is essential to prevent severe

consequences, including significant financial and commercial losses. Organization like General Data Protection Regulationwas established during 2018withrecognized as standard aimed at ensuring user data privacy, encouraging AI and machine learning frameworks to adhere to these regulations [4].

Machine learning and AI models require substantial data for training to produce refine and better outputs. These models require user data to develop precise predictive capabilities, ensuring user privacy remains paramount. Some organizations must exchange data for collaborative efforts that enhance performance while safeguarding data privacy and confidentiality [5]. On edge devices, such as mobile phones, users interact with numerous applications daily, exposing a significant volume of private data. To handle the challenge of using data for training while maintaining privacy, Federated Learning has appear as a answer which is introduced by McMahan [6], Federated Learning is collaborative machine learning method able to trained models available at edge devices instead of a centralize server. In Federated Learning (FL), data is not transferred between devices; rather, only model updates (gradient updates) is transfer between server. The server merges available updates for creation of universal model without accessing raw data from edge devices, thereby maximizing user data privacy. Federated Learning has emerged in reply to the growing demand for data privacy, providing a safer alternative to centralized machine learning methods and reducing related costs [7].

## 3. The Problem

Accurate risk forecasting is essential for maintaining stability and making sound financial decisions. Traditional machine learning models require localized data collection, which raises serious privacy concerns especially given the sensitivity of financial data[8][9], Federated Learning (FL) offers a circuitous approach, and it enables models to be trained different multiple organizations in absence of raw data [10]. But despite the promise, government studies are not inherently privacy-protective and can still leak important information through the model-update slope.

This research seeks to tackle the crucial challenge of designing a privacy-preserving federated learning algorithm specifically designed to train financial risk prediction models. The key problems to be tackled include:

1. **Privacy Leakage:** Ensuring that the federated learning process does not expose sensitive financial data, either directly or indirectly, through model updates.

2. **Model Accuracy:** Maintaining high predictive accuracy while incorporating privacy-preserving mechanisms.

3. **Communication Overhead:** Minimizing the communication cost between participating institutions, this can be substantial in federated learning setups.

4. **Scalability:** Ensuring the presented algorithm scales efficiently as participation institutions and the volume of data increase.

5. **Regulatory Compliance:** Aligning the developed algorithm with stringent financial regulations and data protection laws, such as GDPR and CCPA.

The research aims to contribute a robust solution that balances privacy, accuracy, and efficiency, facilitating the safe and effective use of federated learning in financial risk prediction.

## 4. Gap in the body of knowledge and motivation of the research

### 4.1. Data Privacy in Federated Learning

While federated learning (FL) inherently provides structure for learning of models without sharing raw data, there remain significant challenges in ensuring robust data privacy and security [11][12]. Existing FL methods often expose vulnerabilities to potential data breaches and inference attacks, which can jeopardize the privacy of confidential financial data [13]. Current solutions need further refinement and integration to effectively safeguard against these threats without severely impacting model performance.

### 4.2. Algorithm Efficiency and Scalability

The implementation of federated learning in financial risk prediction models faces issues related to algorithm efficiency and scalability [14][15]. The computational and communication overhead involved in training complex models over distributed and diverse data sources can be significant. There is a need for developing more efficient FL algorithms that can handle large-scale financial datasets, reduce latency, and manage resource constraints while maintaining high prediction

accuracy [16]. Addressing this gap involves optimizing communication protocols, model update mechanisms, and resource allocation strategies.

## 5. Objective and research query

It helps in develop federated learning algorithm that preserves privacy specifically designed for training financial risk prediction models. This algorithm will enable financial institutions to collaboratively build robust and accurate predictive models while ensuring data privacy, regulatory compliance, and ethical data handling.

1. What are the current financial risk prediction models for ensuring data privacy?
2. What are the challenges, limitation and performance implications of incorporating privacy-preserving techniques into financial risk prediction?
3. What are the common types of attacks that target financial risk model?
4. What are most effective model/algorithm for securing privacy?
5. How can effectiveness of these algorithms be evaluated?

By highlighting these research questions, this study aims to offer a thorough exploration of designing, implementing, and assessing the impact of a federated learning algorithm focused on preserving privacy in financial risk prediction. This endeavor ultimately contributes to advancements in both machine learning and the protection of financial data privacy.

## 6. Objectives of the study

1. Analyze the challenges, limitations, and performance implications of challenges into various systems.
2. Propose a robust model and algorithm for detecting anomalies in financial transactions, focusing on accuracy and efficiency.
3. Assess the efficacy of the proposed algorithms in maintaining privacy within context of federated learning.

## 7. Theoretical background

Federated Machine Learning is a distributed structure which helps in training of models across various decentralized points, each storing its own data locally. This method avoids the



Kumar

requirement of transfer sensitive data to a centralize server, promoting enhanced privacy, data security, and computational efficiency. By allowing the model to get training on local datasets, FML reduces data movement, improving scalability and minimizing potential risks associated with data centralization.

### 7.1 Key Concepts

a. Decentralized Data Processing: Conventional machine learning methods are depend on central data repositories for training models, Federated Machine Learning (FML) keeps data on individual devices where each device trains autonomously using its local data and sends only model updates parameters with a central server.

b. Collaborative Training: Centralize server combine model updates from various devices to build a universal model. This process is iterative, with each round involving local training on devices followed by central aggregation. The cycle continues until the model converges, gradually improving as more updates are integrated from multiple devices.

c. Privacy Preservation: FML improves privacy by holding sensitive data on individual devices rather than sending to the central server, only parameters to the model are communicated which helps in minimizes the risk of revealing personal information.

### 7.2 Training Process

Federated Learning is distributed method for learning which permit different devices and servers to work collaborative way to train a model with no requirement to send their local data. This method promotes privacy and security by holding data on the device while utilizing the advantages of a diverse and distributed data pool [19].

In Federated Learning, the learning begins with a centralize server that initializes a model. This initial model is sent out to each participating devices, where each node independently trains the model using its own data [20]. It is crucial that data resides on the client device, ensuring that privacy is secure as illustrated in figure 2.

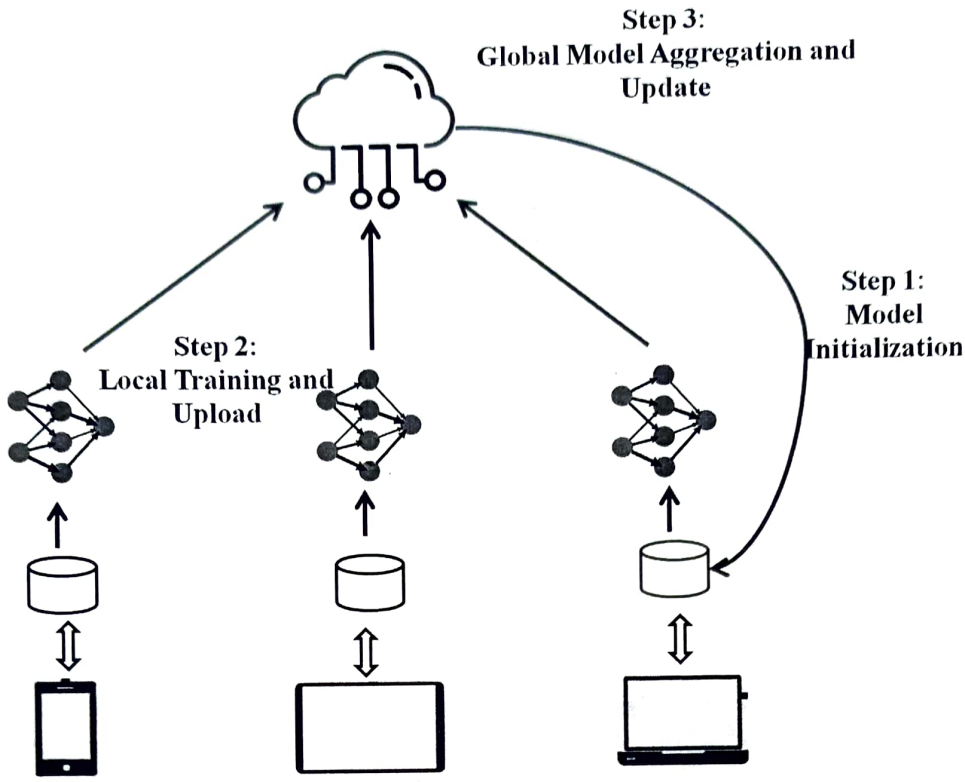


Fig 2: Federated Learning Process

After the localize training is done, every client updates its model parameters on its own dataset. These updates are reverting back to the centralize server where it combine the parameters from all clients to form a new universal model [21]. A commonly used technique is Federated Averaging (FedAvg), which add together a weighted average of the local model updates. To enhance privacy, secure aggregation methods will be applied to guarantee that server only receives the consolidated update and not the individual updates. Once aggregation is finished, central server updates universal model with these aggregated parameters. The updated model is subsequently broadcasted return to the clients, replacing previous version. This cycle of local training, uploading updates, aggregation, and broadcasting is repeated over multiple rounds. The process pursuetill a predefined convergence criterion is achieved, such as achieving target accuracy, completing a specified number of iterations and observing a plateau in model efficiency.

### 7.3 Federated Learning Life Cycle

The life cycle of Federated Learning (FL) encompasses several stages, from the initial setup to the deployment and monitoring of the trained model. This life cycle ensures that the FL process is



Saurav Kumar

systematic, secure, and efficient. The process begins with problem definition and feasibility analysis, where the machine learning problem is identified, and the suitability of FL is evaluated by considering factors like data distribution, privacy requirements, computational resources, and network infrastructure.

After data preparation and device selection take place. This stage ensures that the data on each participating device is preprocessed and ready for training, including tasks like data cleaning, normalization, and augmentation. Simultaneously, suitable edge devices are selected based on their availability, computational power, and data relevance.

Following this, initial model training and distribution occur. A global model is initialized, either pre-trained or from scratch, and then distributed to all selected edge devices to begin the training process. Each device then engages in local training, where each client refresh the global model with its local data through several epochs of training and computes model updates (e.g., gradients or weight changes).

These locally generated model updates are then communicated back to the central server, confirming only updates are shared to maintain privacy. Techniques like model compression and reducing update frequency are used to minimize communication overhead. Subsequently, the central server consolidates updates from all participating devices to produce a new universal model, often using Federated Averaging (FedAvg) or secure aggregation protocols. The aggregated model is updated to incorporate the changes from all devices.

Multiple training cycles are performed where devices train a model locally, share updates, and then those updates are combined to create a better overall model. This cycle continues until the overall model performs well enough, reaching a desired level of accuracy. Once convergence is achieved, the model undergoes evaluation and validation using a separate validation dataset to ensure it meets the desired performance criteria. Different evaluation metrics track how well the model is doing by using different parameters for evaluation.

After training global model is deployed to the production environment for inference on new data. Continuous monitoring is essential to ensure the deployed model remains effective and to identify any potential issues or drifts in data distribution. Regular updates to the global model are necessary, achieved through repeating the FL process to incorporate new data and address changes in the underlying data distribution. A feedback loop is implemented to collect performance data and user feedback, which can be used to further refine the model.



Gaurav Kumar

The Federated Learning life cycle encompasses a comprehensive process that oversees the development, training, deployment, and maintenance of machine learning models in a disperse manner. By following these stages, organizations can leverage distributed data sources while maintaining high standards of privacy and security, ultimately creating robust and effective models tailored to their specific needs.

### 7.4 Privacy Mechanisms in Federated Learning

In Federated Learning, various privacy techniques have been employed to safeguard sensitive user data while allowing effective model training and collaboration across decentralized devices. These privacy techniques can be distinguish into several categories according to their methodologies. Some types of privacy mechanisms used in FL:

**a. Differential Privacy (DP):** It is technique designed to quantify and minimize privacy risks when processing data. It involves adding noise to model updates or aggregates. This noise ensures that an adversary cannot deduce sensitive information about any individual data point, regardless of its inclusion or exclusion in the dataset. By making the impact of any single data sample negligible, DP helps safeguard privacy while still enabling effective model training.

**b. Secure Aggregation:** It enable multiple edge devices to collaboratively compute aggregated model updates without revealing contributions. Such techniques allow devices to perform computations on encrypted data, eliminating the requirement for decryption and protecting the information of both raw data and intermediate model updates. These methods ensure that sensitive information remains secure during the aggregation phase, preventing unauthorized access to user data or model parameters.

c. **Federated Averaging with Differential Privacy**: It is a popular algorithm helps in aggregating model updates in FL. By integrating Differential Privacy into FedAvg, noise is added to the gradients before aggregation, ensuring that individual contributions are privacy-preserving. FedAvg with Differential Privacy attain a equilibrium between model accuracy and privacy protection, enabling Federated Learning (FL) models to achieve competitive performance making security of data.

d. **Local Differential Privacy**: Local Differential Privacy mechanisms operate directly on edge devices to perturb user data before model training. Instead of adding noise during aggregation, Local Differential Privacy adds noise samples prior to they are utilized for training, safeguarding sensitive information at its origin and minimizing the potential for privacy breaches throughout the Federated Learning (FL) process.Local Differential Privacy enhances user privacy by preserving the confidentiality of raw data on individual devices.

e. **Model Personalization**: Model Personalization techniques enable FL models to be tailored to individual users' preferences and characteristics without exposing raw data. By incorporating user-specific updates during model aggregation, FL models can adapt to diverse data distributions while preserving user privacy. Model Personalization allows FL systems to provide personalized services without compromising data confidentiality.

These privacy mechanisms are crucial for fostering trust, ensuring compliance, and gaining user acceptance in Federated Learning (FL) applications, especially in sensitive domains like healthcare, finance, and personal assistance. By incorporating these mechanisms into FL frameworks and protocols, organizations can uphold stringent privacy standards while leveraging the benefits of decentralized data processing and collaborative model training.

### 7.5 LSTM for Anomaly Detection

Long Short-Term Memory networks are efficient for sequential data due to their capability to capture long-term dependencies. In anomaly detection, primary goal is to detect deviations from normal sequential patterns.

- Sequential Nature: LSTM can be trained on historical transaction data, where each transaction is dependent on the previous ones. It can learn typical transaction patterns such as frequency, amounts, and timing.
- Modeling Normal Behavior: Once the LSTM model is trained, it will learn what normal sequences of transactions look like. For example, if a user typically makes small transactions, an unusually large transaction would appear as an anomaly.
- Anomaly Detection: The LSTM network predicts the next transaction in the sequence. If the predicted value deviates significantly from the actual value, this is flagged as an anomaly. The deviation could be in the form of the transaction amount, timing, frequency, or other features.

LSTM-based Anomaly Detection:

- Training: LSTM is trained on historical data, learning normal transaction patterns.
- Prediction: For each new transaction, the model predicts the expected value (next transaction or sequence).
- Anomaly Scoring: The model compares the predicted value with the actual value. A large difference indicates an anomaly.

### 7.6 Autoencoders for Anomaly Detection

Autoencoders are unsupervised neural networks designed to compress data and then reconstruct it. They are widely used in anomaly detection because they learn to recreate normal data and fail to reconstruct anomalies.

How Autoencoders work for anomaly detection:

- Normal Data Learning: The autoencoder learns to encode and decode normal transaction patterns. It does this by compressing the input data into a encoding and then reconstructing it back to decoding.
- Anomaly Detection: After training, the autoencoder will have a low reconstruction error when processing normal data, as it has learned the pattern. However, when faced with an anomalous or unseen input (e.g., fraudulent transactions), the reconstruction error will be high, signaling an anomaly.

Autoencoder-based Anomaly Detection:

- Training: The autoencoder is trained on normal transaction data (without anomalies).
- Reconstruction: For new data, the autoencoder attempts to reconstruct the input.
- Anomaly Scoring: The reconstruction error is calculated. If the error is larger than a pre-defined threshold, it is flagged as an anomaly..

## 8. Scope of study

This study focuses on developing and assessing a federated learning algorithm that preserves privacy, tailored specifically for training models predicting financial risks [22]. This includes designing a federated learning framework suitable for financial institutions; integrating advanced privacy-preserving methods [23]. Evaluate the performance and privacy effectiveness of the developed models using metrics such as accuracy, convergence rate, computational efficiency,

and privacy preservation. Additionally, it will ensure that the framework complies with data protection regulations like GDPR and CCPA, offering guidance for regulatory compliance. The study will also develop and test secure aggregation methods to combine model updates without compromising data privacy. Furthermore, it will identify and address practical challenges in deploying the algorithm in the financial sector, offering solutions and recommendations for real-world implementation. Finally, the study aims to promote ethical data use by ensuring responsible data handling and fostering trust in data-driven financial services.

## 9. Literature review

Jeon, Beomyeolet al, 2021 [24] introduce a decentralized aggregation approach that eliminates the need for a central aggregator, enhancing both privacy and scalability. By leveraging cryptographic techniques and privacy-preserving mechanisms, the method ensures the confidentiality of individual model updates while maintaining efficiency and accuracy. Experimental results show the effectiveness of preserving privacy without compromising performance of FL models. Overall, the paper offers a significant contribution to advancing FL systems, making them more secure and suitable for real-world deployment.

Mammenet al., 2021 [25], examines the potential benefits of FL, such as enhanced privacy, scalability, and efficiency, in various application domains. Additionally, it discusses the challenges faced in implementing FL systems, including communication overhead, heterogeneous data distributions, and model aggregation complexities. By offering a comprehensive analysis of FL, this research informs both researchers and practitioners about the opportunities and obstacles in deploying FL solutions, guiding future developments in this emerging field.

Liu et al., 2021 [26], advocates for keeping data local to devices and leveraging FL techniques for collaborative model training. By distributing the learning process across edge devices and aggregating model updates instead of raw data, the proposed method enhances privacy while enabling efficient edge computing applications. This research shows a encouraging solution for addressing data privacy concerns in edge computing scenarios, contributing to the advancement of secure and privacy-preserving edge computing systems.

Mo, Fan et al.,2021 [27] introduces approach to privacy-preserving federated learning by leveraging rely to execution environments (TEEs). They propose a framework that employs Trusted Execution Environments (TEEs) to secure model training processes while securing the

privacy of user information. By encapsulating sensitive operations within TEEs, the proposed PPFL scheme ensures that individual user data remains confidential and protected from unauthorized access. This research contributes to advancement in security and privacy of federated learning systems, offering a robust solution that can be deployed across various mobile and distributed computing environments.

Chamikara et al., 2021 [28] Federated learning preserves privacy by keeping data itself and send only model updates, not raw data. However, a major challenge is the high communication cost, particularly when training large neural networks for tasks like financial risk prediction. To address this, methods such as compression and quantization may be used to reduce model updates without compromising the precision of risk prediction models.

Wu et al., 2022 [14], Propose a scheme that adapts privacy protections based on the sensitivity of the data being exchanged. By integrating differential privacy mechanisms into the FL framework, the proposed scheme ensures that individual user data remains confidential during model training. This research enhances the development of federated learning systems that prioritize security and privacy, offering a flexible solution that adapts to varying privacy requirements across different data types and applications.

Song, Jingcheng et al.,2022 [29] Propose a novel method for safeguarding data aggregation in federated learning (FL) while ensuring privacy protection. Their design efficiently aggregates data while maintaining user privacy. By leveraging advanced cryptographic techniques and federated learning methodologies, EPPDA guarantees the security and confidentiality of user data throughout the aggregation process. This scheme presents encouraging solution to handle privacy issue in federated learning, thereby advancing the development of secure and scalable FL systems.

Wang et al., 2022 [30], Federated learning enables distributed machine learning with safeguarding privacy of end devices' data. However, frequent access to cloud servers can cause substantial delays in transmitting model updates. Moreover, local models may sometimes conflict with the global model, leading to heightened communication expenses. Current research primarily targets reducing communication rounds or handling flawed local data, often neglecting the latency caused by intense server concurrency. To tackle this challenge, we investigate an optimization framework based on edge computing. This framework aims to reduce direct connection of end devices to the central server and minimization of transmission of local updates. It involves clustering devices within the identical network area and deploying edge nodes across



Gaurav Kumar

multiple locations to serve as intermediaries between the cloud and end devices. This strategy helps alleviate latency issues stemming from high server concurrency.

Makkar et al., 2022 [31],A significant volume of data is processed across numerous websites, AI tools, and data silos, posing challenges in sharing data due to administrative, organizational, and security constraints. Federated machine learning (FML) addresses this by deploying machine learning algorithms to data sources, training models locally, and aggregating the results while preserving consumer privacy through local data processing. This research pioneers the application of FML in detecting Internet attacks, focusing particularly on identifying malicious spam images that can disrupt AI systems and generate irrelevant search outcomes. Our key contributions include enhancing customer search queries by employing FML to detect image spam. We conducted a comprehensive evaluation of FML's performance across diverse scenarios, including balanced and imbalanced data distributions, scalability considerations, and communication overhead. Our findings highlight FML's efficacy in practical applications, notably in distinguishing between legitimate images and spam in advertisements, thereby enhancing search relevance. The evaluated results underscore the effectiveness of the advanced FedLearnSP approach in image spam detection.

Nguyen et al., 2023 [32], Propose strategies to strengthen privacy and security of federated learning systems, ensuring protection of sensitive data throughout model training and aggregation processes. By leveraging cryptographic techniques, differential privacy, and secure aggregation protocols, the paper offers solutions to mitigate privacy risks and safeguard against adversarial attacks in FL environments. This research enhances the development of robust and privacy-focused FL frameworks, promoting trust and reliability in distributed machine learning systems.In the view of financial risk prediction, use of machine learning algorithms has gained significant attention due to their potential in improving the accuracy and efficiency of risk assessment(Zhao et al.,2023) [33]. To safeguard the privacy of model updates in federated learning, secure multi-party computation techniques can be employed. They introduce Cecilia, a federated architecture integrating cloud, edge, and client components, along with a novel algorithm named ACFL. These innovations target communication bottlenecks in federated learning, aiming to enhance network throughput and overcome challenges posed by unreliable networks and fluctuating bandwidth. They develop a algorithm for securing privacy with federated learning specifically tailored for training financial risk prediction models (Lim et al.,2023)[34].

H Wang et al, 2024 [35] they introduces a Privacy-securing Federated Learning method with Partial Low-Quality Data. It enhances privacy and security using a distributed Paillier cryptographic mechanism and improves model quality by evaluating and minimizing low-quality data's impact. Results on the MNIST dataset show that PPFL-LQDP effectively trains models while protecting data privacy and security, outperforming related schemes.

Li et al, 2024 [36] shows that rapid growth of Smart Internet of Things generates vast volume of data, making traditional machine learning inadequate for training complex models. Federated learning approach for training models in decentralized networks, addresses with large, diverse datasets while ensuring data privacy. Edge computing enhances this by processing data near its source, reducing latency, and providing high-bandwidth communication and stable networks, easing the burden on cloud centers. Combining edge computing with federated learning optimizes computing and data for the edge device. They examine the expansion of federated learning, its principles and security measures like cryptographic technologies, differential privacy, and adversarial training. It also discusses issue and future research for integrating edge IoT with federated learning.

## 10. Research Methodology

The research methodology for developing a privacy-preserving federated learning algorithm for training financial risk prediction models involves a structured approach to address the complexities of privacy, regulatory compliance, and performance in the financial sector as shown in figure 3.

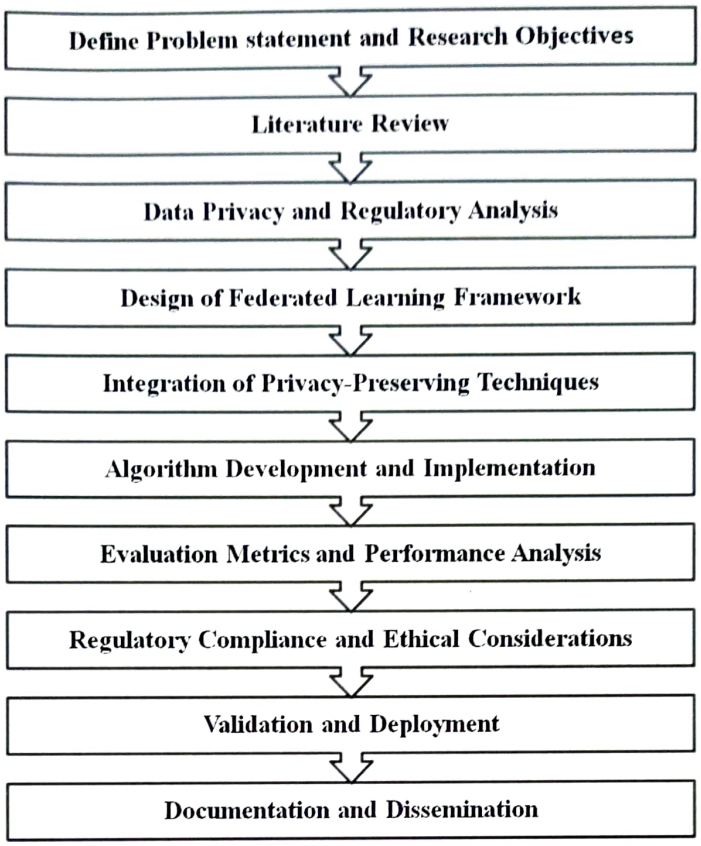
Define Problem statement and Research Objectives
Literature Review
Data Privacy and Regulatory Analysis
Design of Federated Learning Framework
Integration of Privacy-Preserving Techniques
Algorithm Development and Implementation
Evaluation Metrics and Performance Analysis
Regulatory Compliance and Ethical Considerations
Validation and Deployment
Documentation and Dissemination

Fig 3: Research Methodology for developing a privacy-preserving federated learning algorithm

i. **Define Problem statement and Research Objectives**: Define the research problem and articulate the specific objectives of the study. Outline the scope of the research, including the types of financial risk prediction models to be addressed.

ii. **Literature Review**: Perform an extensive survey of the current body of literature concerning federated learning (FL), privacy-preserving techniques, and their applications in financial risk prediction. Identify gaps and limitations in current methodologies and frameworks that the research aims to address.

iii. **Data Privacy and Regulatory Analysis**: Analyze regulatory requirements relevant to data privacy in financial sectors (e.g., GDPR, CCPA, financial industry standards).Define how these regulations impact the design and implementation of the federated learning algorithm.

iv. **Creation of a Federated Learning Framework:** Establish a conceptual framework for federated learning that meets the specific requirements of financial institutions. Additionally, outline the responsibilities of central servers and client nodes, communication protocols, and techniques for the secure aggregation of model updates.

v. **Adoption of Privacy-Enhancing Techniques:** Embed advanced privacy-preserving approaches such as differential privacy, secure multi-party computation, and homomorphic encryption within the federated learning architecture to safeguard sensitive financial data.

vi. **Creation and Execution of Algorithms:** Develop the federated learning algorithm in line with the suggested framework and the incorporated privacy-preserving techniques.

vii. **Assessment Metrics and Performance Evaluation:** Specify performance indicators such as model accuracy, convergence rate, and computational efficiency. Conduct experiments to evaluate the effectiveness of the privacy-preserving federated learning algorithm compared to centralized methods.

viii. **Compliance with Regulations and Ethical Considerations:** Evaluate the algorithm's adherence to data protection legislation and the ethical issues regarding data utilization and transparency. Address potential ethical dilemmas in the implementation of machine learning models within financial contexts.

ix. **Validation and Deployment**: Validate the effectiveness and practicality of the developed algorithm through testing and validation with synthetic or real-world financial datasets. Provide recommendations for deploying the privacy-preserving federated learning algorithm in real-world financial environments.

x. **Documentation and Dissemination**: Document research findings, methodologies, and outcomes in academic papers, technical reports, and presentations. Share insights and lessons gained to contribute to progress of privacy-preserving federated learning for financial risk prediction.

## 11. Discussion

Federated Learning (FL) acting an innovative method to collaborative machine learning, offering decentralized model training while making the privacy of user information. However, distributed nature of it introduces vulnerabilities which malicious actors can exploit through different privacy attacks [37]. Understanding and addressing these attacks is paramount for maintaining the integrity and confidentiality of FL systems. Here's a discussion on common FL privacy attacks and their implications [38]:

Privacy attacks in federated learning encompass a range of techniques aimed at compromising confidentiality and integrity of user. One prevalent attack is model inversion, where adversaries attempt to reverse-engineer sensitive training data from the global model or its gradients. By

analyzing model outputs or gradients, attackers can infer details about individual training samples, potentially revealing personal or proprietary information [12]. This type of attack undermines the fundamental privacy guarantees of FL systems, exposing users to privacy breaches and compromising sensitive data.

Another significant threat arises from membership inference attacks, where adversaries attempt to ascertain whether specific data points were included in the training dataset [39]. By exploiting information leakage in the model's outputs, attackers can infer the presence or absence of individual training samples, compromising user privacy and confidentiality [40]. This knowledge can be leveraged for targeted profiling or discriminatory practices, posing ethical and legal concerns for FL deployments.

Model extraction attacks pose a severe risk to FL systems by allowing adversaries to steal proprietary model architectures and intellectual property. Through careful analysis of the model's responses to queries, attackers can construct surrogate models that closely mimic the original, potentially revealing subtle information about the training data. Such attacks undermine competitive advantage of FL participants and may lead to economic losses and reputational damage.

Furthermore, model poisoning attacks pose a substantial threat to integrity and reliability of federated learning (FL) models. Adversaries achieve this by injecting malicious data or perturbations during the training process, with the aim of manipulating model's behavior. leading to erroneous forcast, biased outcomes, or security breaches. These attacks can have profound implications, especially in safety-critical applications like healthcare or autonomous vehicles, where the reliability of ML models is crucial.

Lastly, differential privacy attacks exploit vulnerabilities in the mechanisms used to achieve privacy guarantees in FL. Adversaries might analyze noisy model updates or aggregates to infer subtle information about individual data samples, compromising the effectiveness of privacy-preserving techniques and violating regulatory requirements.

To mitigate these privacy attacks, FL systems must employ robust privacy-preserving mechanisms, secure communication protocols, and adversarial defense techniques. Moreover, ongoing research and collaboration are essential for developing resilient FL frameworks that withstand evolving privacy threats and maintain user trust in real-world deployments. Only by addressing these challenges it can achieve its full capacity as a privacy-preserving and effective machine learning across various domains and applications.



Jaiswal Kumar

## 12. Conclusions and Future Scope

Privacy-preserving Federated Learning (FL) has become a key approach in modern machine learning frameworks, enabling joint training among distributed devices while maintaining user privacy [41]. Leveraging techniques like Differential Privacy, Secure Aggregation, and Local Differential Privacy, FL systems have made significant strides in mitigating privacy risks and safeguarding sensitive data [42]. Nonetheless, challenges persist in combating privacy attacks and fortifying the resilience of privacy-preserving mechanisms in FL.

There's a compelling need for ongoing exploration into advanced privacy techniques. Techniques like homomorphism encryption and multi-party computation hold potential in bolstering the privacy assurances of FL systems, countering emerging privacy threats with robust solutions.

The design of privacy-aware model architectures presents an exciting frontier. Models engineered with inherent privacy preservation during training could significantly mitigate the risks of privacy breaches, fostering transparent and accountable FL systems.

Moreover, compliance with evolving privacy regulations, like GDPR, remains paramount. Establishing standardization frameworks and best practices can streamline development processes and foster consistent privacy protection across FL deployments.

Empowering users with greater control over their data and privacy preferences is also essential. Fine-grained consent mechanisms and transparent data usage policies can enhance user trust and engagement in collaborative machine learning initiatives.

Real-world deployments across diverse domains will provide invaluable insights into practical challenges and opportunities. Large-scale pilot studies in healthcare, finance, telecommunications, and manufacturing will inform the refinement of privacy-preserving FL methodologies for real-world applications.

The future of privacy-preserving FL holds promise in revolutionizing collaborative machine learning while safeguarding user privacy rights and ethical principles [43]. By embracing interdisciplinary collaboration and continued research endeavors, FL stands poised to redefine how machine learning models are trained and deployed in privacy-conscious environment, fostering a more inclusive, transparent, and trustworthy digital landscape.

## 13. References

[1] Wu, Yuncheng, Shaofeng Cai, Xiaokui Xiao, Gang Chen, and Beng Chin Ooi. *"Privacy preserving vertical federated learning for tree-based models." arXiv preprint arXiv:2008.06170* (2020).

[2] Kairouz, P., McMahan, H. B., Avent, B., Bellet, A., Bennis, M., Bhagoji, A. N., & Suresh, A. T. (2019). *Advances and Open Problems in Federated Learning.* arXiv preprint arXiv:1912.04977.

[3] Zhang, Chen, Yu Xie, Hang Bai, Bin Yu, Weihong Li, and Yuan Gao. *"A survey on federated learning."Knowledge-Based Systems* 216 (2021): 106775.

[4] Truong, Nguyen, Kai Sun, Siyao Wang, Florian Guitton, and YiKe Guo. *"Privacy preservation in federated learning: An insightful survey from the GDPR perspective."* *Computers & Security* 110 (2021): 102402.

[5] Mohri, Mehryar, Gary Sivek, and Ananda Theertha Suresh. *"Agnostic federated learning."* In *International Conference on Machine Learning*, pp. 4615-4625. PMLR, 2019.

[6] McMahan, H. B., Moore, E., Ramage, D., Hampson, S., & y Arcas, B. A. (2017). *Communication-Efficient Learning of Deep Networks from Decentralized Data.* Proceedings of the 20th International Conference on Artificial Intelligence and Statistics (AISsTATS).

[7] Bonawitz, K., Eichner, H., Grieskamp, W., Huba, D., Ingerman, A., Ivanov, van Overveldt, T. (2019). *Towards Federated Learning at Scale: System Design.* Proceedings of the 2nd SysML Conference.

[8] Kairouz, Peter, H. Brendan McMahan, Brendan Avent, Aurélien Bellet, Mehdi Bennis, Arjun Nitin Bhagoji, Kallista Bonawitz et al. "*Advances and open problems in federated learning.*" *Foundations and trends® in machine learning* 14, no. 1–2 (2021): 1-210.

[9] Ferrag M.A., Friha O., Hamouda D., Maglaras L., Janicke H.Edge-IIoTset: *A new comprehensive realistic cyber security dataset of IoT and IIoT applications for centralized and federated learning*, IEEE Access, 10 (2022), pp. 40281-40306.

[10] Li, Qinbin, Bingsheng He, and Dawn Song. "*Model-contrastive federated learning.*" In Proceedings of the IEEE/CVF conference on computer vision and pattern recognition, pp. 10713-10722. 2021.

[11] Luo, Bing, Xiang Li, Shiqiang Wang, Jianwei Huang, and Leandros Tassiulas. *"Cost-effective federated learning design."* In IEEE INFOCOM 2021-IEEE Conference on Computer Communications, pp. 1-10. IEEE, 2021.

[12] Xia, Q., Ye, W., Tao, Z., Wu, J., & Li, Q. (2021). "*A survey of federated learning for edge computing: Research problems and solutions*". High-Confidence Computing, 1(1), 100008. https://doi.org/https://doi.org/10.1016/j.hcc.2021.100008

[13] Briggs, Christopher, Zhong Fan, and Peter Andras. *"A review of privacy-preserving federated learning for the Internet-of-Things." Federated Learning Systems: Towards Next-Generation AI* (2021): 21-50.

[14] Ma, Zhuoran, Jianfeng Ma, Yinbin Miao, Yingjiu Li, and Robert H. Deng. *"ShieldFL: Mitigating model poisoning attacks in privacy-preserving federated learning."* IEEE Transactions on Information Forensics and Security17 (2022): 1639-1654.

[15] Yang, Qiang. "*Toward responsible ai: An overview of federated learning for user-centered privacy-preserving computing*." ACM Transactions on Interactive Intelligent Systems (TiiS)11, no. 3-4 (2021): 1-22.

[16] Xia, Q., Ye, W., Tao, Z., Wu, J., & Li, Q. (2021*). "A survey of federated learning for edge computing: Research problems and solutions"*.High-Confidence Computing, 1(1), 100008. https://doi.org/https://doi.org/10.1016/j.hcc.2021.100008.

[17] Wu, Xiang, Yongting Zhang, Minyu Shi, Pei Li, Ruirui Li, and Neal N. Xiong. *"An adaptive federated learning scheme with differential privacy preserving."* Future Generation Computer Systems 127 (2022): 362-372.

[18] Nguyen J., Malik K., Zhan H., Yousefpour A., Rabbat M., Malek M., Huba D., *Federated learning with buffered asynchronous aggregation*International Conference on Artificial Intelligence and Statistics, PMLR (2022), pp. 3581-3607.

[19] Brecko, A., Kajati, E., Koziorek, J., &Zolotova, I. (2022). *Federated Learning for Edge Computing: A Survey*. Applied Sciences, 12(18).https://doi.org/10.3390/app12189124

[20] Ferrag M.A., Friha O., Maglaras L., Janicke H., Shu L.*Federated deep learning for cyber security in the Internet of Things: Concepts, applications, and experimental analysis*, IEEE Access, 9 (2021), pp. 138509-138542.

[21] Chen Z., Liao W., Hua K., Lu C., Yu W.*Towards asynchronous federated learning for heterogeneous edge-powered Internet of Things,* Digit. Commun. Netw., 7 (3) (2021), pp. 317-326.

[22] Lim, W. Y. B., Luong, N. C., Hoang, D. T., Jiao, Y., Liang, Y.-C., Yang, Q., Niyato, D., & Miao, C. (2020). *"Federated Learning in Mobile Edge Networks: A Comprehensive Survey"*.

IEEE Communications Surveys & Tutorials, 22(3), 2031–2063.https://doi.org/10.1109/COMST.2020.2986024.

[23] Li, D., Han, D., Weng, TH.*Blockchain for federated learning toward secure distributed machine learning systems: a systemic survey*.Soft Comput 26, 4423–4440 (2022). https://doi.org/10.1007/s00500-021-06496-5.

[24] Jeon, Beomyeol, S. M. Ferdous, Muntasir Raihan Rahman, and Anwar Walid. "*Privacy-preserving decentralized aggregation for federated learning*." In IEEE INFOCOM 2021-IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp. 1-6. IEEE, 2021.

[25] Mammen, Priyanka Mary. "*Federated learning: Opportunities and challenges*." *arXiv preprint arXiv:2101.05428* (2021).

[26] Liu, Gaoyang, Chen Wang, Xiaoqiang Ma, and Yang Yang. "*Keep your data locally: Federated-learning-based data privacy preservation in edge computing*." *IEEE Network* 35, no. 2 (2021): 60-66.

[27] Mo, Fan, Hamed Haddadi, KleomenisKatevas, Eduard Marin, Diego Perino, and Nicolas Kourtellis. "*PPFL: privacy-preserving federated learning with trusted execution environments*." In Proceedings of the 19th annual international conference on mobile systems, applications, and services, pp. 94-108. 2021.

[28]Chamikara, Mahawaga Arachchige Pathum, Peter Bertok, Ibrahim Khalil, Dongxi Liu, and Seyit Camtepe. "*Privacy preserving distributed machine learning with federated learning*." *Computer Communications* 171 (2021): 112-125.

[29] Song, Jingcheng, Weizheng Wang, Thippa Reddy Gadekallu, Jianyu Cao, and Yining Liu. "*Eppda: An efficient privacy-preserving data aggregation federated learning scheme*." IEEE Transactions on Network Science and Engineering(2022).

[30] T. Wang, Y. Liu, X. Zheng, H. -N. Dai, W. Jia, and M. Xie, *"Edge-Based Communication Optimization for Distributed Federated Learning,"* in IEEE Transactions on Network Science and Engineering, vol. 9, no. 4, pp. 2015-2024, 1 July-Aug. 2022, doi: 10.1109/TNSE.2021.3083263.

[31] A. Makkar, U. Ghosh, D. B. Rawat and J. H. Abawajy, *"FedLearnSP: Preserving Privacy and Security Using Federated Learning and Edge Computing,"* in IEEE Consumer Electronics Magazine, vol. 11, no. 2, pp. 21-27, 1 March 2022, doi: 10.1109/MCE.2020.3048926.

[32] Nguyen, Truc, and My T. Thai. "*Preserving privacy and security in federated learning."IEEE/ACM Transactions on Networking* (2023).

[33] Zhao, Bin, Kai Fan, Kan Yang, Zilong Wang, Hui Li, and Yintang Yang. *"Anonymous and privacy-preserving federated learning with industrial big data."* IEEE Transactions on Industrial Informatics 17, no. 9 (2021): 6314-6323.

[34] W. Y. B. Lim., "*Decentralized Edge Intelligence: A Dynamic Resource Allocation Framework for Hierarchical Federated Learning*," inIEEE Transactions on Parallel and Distributed Systems, vol. 33, no. 3, pp. 536-550, 1 March 2022, doi: 10.1109/TPDS.2021.3096076.

[35] Wang, H., Wang, Q., Ding, Y. *"Privacy-preserving federated learning based on partial low-quality data". J Cloud Comp***13**, 62 (2024). https://doi.org/10.1186/s13677-024-00618-8.

[36] Li, H., Ge, L. & Tian, L. *"Survey: federated learning data security and privacy-preserving in edge-Internet of Things". ArtifIntell Rev***57**, 130 (2024). https://doi.org/10.1007/s10462-024-10774-7.

[37] Lyu, Lingjuan, Han Yu, and Qiang Yang. "*Threats to federated learning: A survey.*" *arXiv preprint arXiv:2003.02133* (2020).

[38] Liu, Y., Zhang, L., Cheng, L., & Liu, J. (2020). *Secure and Efficient Federated Transfer Learning*. IEEE Transactions on Computers, 70(6), 1002-1015.

[39] SameeraK.M., SerenaNicolazzo,MarcoArazzi, AntoninoNocera, RafidhaRehimanK.A., VinodP., MauroConti , "*Privacy-preserving in Blockchain-based Federated Learning systems*" ,Computer Communications, Volume 222, 1 June 2024, Pages 38-67.

[40] Duan, Q.; Hu, S.; Deng, R.; Lu, Z. *Combined Federated and Split Learning in Edge Computing for Ubiquitous Intelligence in Internet of Things: State-of-the-Art and Future Directions.* Sensors 2022, 22, 5983. https://doi.org/10.3390/s22165983

[41] Wang, S., Tuor, T., Salonidis, T., Leung, K. K., Makaya, C., He, T., & Chan, K. (2019). *Adaptive Federated Learning in Resource Constrained Edge Computing Systems*. IEEE Journal on Selected Areas in Communications, 37(6), 1205-1221.

[42] Li, T., Sahu, A. K., Talwalkar, A., & Smith, V. (2020). *Federated Learning: Challenges, Methods, and Future Directions.* IEEE Signal Processing Magazine, 37(3), 50-60.

[43] Fang, Chen, Yuanbo Guo, Yongjin Hu, Bowen Ma, Li Feng, and Anqi Yin. "*Privacy-preserving and communication-efficient federated learning in internet of things.*"*Computers & Security* 103 (2021): 102199.